2014

EVALUATION FRAMEWORK FOR 3D COLLABORATIVE VIRTUAL ENVIRONMENTS (THE CORE)

Nader Hanna
Macquarie University, nader.hanna@mq.edu.au

Deborah Richards
Macquarie University, deborah.richards@mq.edu.au

EVALUATION FRAMEWORK FOR 3D COLLABORATIVE VIRTUAL ENVIRONMENTS (THE CORE)

Nader Hanna, Computing Department, Macquarie University, Sydney, NSW, Australia, nader.hanna@mq.edu.au

Deborah Richards, Computing Department, Macquarie University, Sydney, NSW, Australia, deborah.richards@mq.edu.au

Abstract

As a consequence of the increasing interest in Collaborative Virtual Environments (CVEs) and the complex nature of such systems, the need for a thorough evaluation framework has become necessary. The current evaluation frameworks suffer from limitations in assessment because they are either confined to evaluation of a specific type of CVE or they focus on a restricted aspect of CVE. This paper presents a framework for THorough Evaluation of COllaborative viRtual Environment (THE CORE). The proposed framework is structured in a four-layered architecture to assure evaluation of the multi-faceted aspects comprising a CVE. The layers evaluate the application usability, tool usability, companion interaction and collaboration outcome. Within each layer, key evaluation tools and factors are provided. In addition, the framework is designed to be generic to be suitable for different types of CVEs. In order to validate the proposed framework, a case study was conducted involving development and evaluation of a 3D CVE. The case study found the four-layered framework to be useful for evaluating both the technical and behavioural aspects of the 3D CVE. As future work, the generality of the framework will be further tested on different types of CVE, potentially leading to modifications and extensions.

Keywords: Collaborative Virtual Environment, Evaluation Framework, Human-Agent Interaction.

1 INTRODUCTION

A Collaborative Virtual Environment (CVE) has been defined as “a computer-based, distributed, virtual space or set of places. In such places, people can meet and interact with others, with agents or with virtual objects.” P.5 (Snowdon et al. 2001). CVEs have been used as a mediation tool to facilitate the human-human collaboration across disparate spaces. Moreover, the concept of CVE includes the collaboration between human participant and virtual entities such as Intelligent Virtual Agents (IVAs).

The past two decades have witnessed increasing interest in using CVEs to facilitate virtual teamwork (Greenberg and Buxton 2008). CVEs have been used in multiple of fields depending on the purpose of use, such as, business (Bishop and Stock 2010) , entertainment, learning (Giraldo et al. 2007) (Lorenzo et al. 2012), training (Holmberg et al. 2006), medicine (Chee and Hooi 2002) and dancing (Zhenyu et al. 2006). CVE has gained a lot of interest due to the evolving growth in networking and telecommunication technologies. Following the increasing interest in CVEs, a vast body of literature work has been published to report the advantages and disadvantages of different CVE designs.

As a consequence of the increasing use of CVEs, there is a need to provide thorough evaluation, particularly when used for serious games such as health, education and training. In his three-layered educational evaluation model, Mikovec et al. (2007) stressed the importance of evaluating a CVE application as it gives an insight into the technical challenges users may face. Evaluation via manual monitoring of CVEs has been found to be error-prone, tedious, and highly impractical especially with large-scale systems (Caballé et al. 2011). The designs of CVEs include many factors that make the evaluation process a difficult task. A large number of factors have been identified that contribute to seamless work and collaboration in VEs, such as immersion, presence, copresence, motion, task involvement, performance, social interaction, coordination and group processes. Moreover, the evaluation of CVE has to take into account several factors like the interaction between the CVE the application domain (Bowman et al. 1999).

Although CVE evaluation has drawn the attention of a number of research projects (Heldal et al. 2007) (Tsiatsos and Andreas 2010), most of the previous studies either use an evaluation method for a single user in the virtual environment (VE) (Hrimech and Merienne 2010), evaluate an exact type of CVE (Tsiatsos and Andreas 2010) such as an educational CVE (Reis et al. 2013) (Wang 2009), or focus on either the technical aspects of a CVE (Oliveira et al. 1999) or human aspects (Whalen et al. 2003) (Wang 2009). To overcome this gap in CVE evaluation, we proposed an evaluation framework that is designed to be sufficiently generic to be able to assess the properties of different CVEs, consider the dual nature of CVEs, that is, both their technical and behavioral aspects. The technical aspect of CVEs covers the compatibility and usability of CE as a software application, while the behavioral aspect considers how participants utilize the CVE user interface to achieve the collaborative task.

The rest of this paper is organized as follows. Section 2 surveys related work on the evaluation of CVEs. Section 3 outlines the proposed evaluation framework along with justification of the design of this framework. A case study to validate the proposed framework (THE CORE) is provided in section 4. Section 5 presents the result of a case study to validate the proposed framework. Finally, in Section 6 some conclusions are drawn and indications given for future lines of research.

2 LITERATURE REVIEW

A number of research papers aimed to evaluate a CVE. Nevertheless, these papers evaluated either a single type of CVE or focused on one or limited aspects of CVE, i.e. technical or behavioural. For example, Wang (2009) evaluated an educational CVE (one type) based on how the students felt about the collaborative learning environment (one aspect). Conducting a wide survey over 180 studies, Tsiatsos and Andreas (2010) compared evaluation frameworks in education and learning CVE. Besides being restricted to the assessment of pedagogical CVEs, the surveyed evaluation frameworks were confined to specific aspects of CVE such as usability, flexibility, pedagogical outcomes, etc. As a result of the survey, Tsiatsos and Andreas (2010) presented a hybrid evaluation framework to collaborative educational VE that combines the technical and pedagogical sides of the educational CVE. However, this evaluation framework focused on the assessment of learning outcome and does not offer a general-purpose evaluation framework. Spada et al. (2005) developed a rating system based on nine characteristic dimensions of collaborative processes. The proposed dimensions did not include an evaluation of the aspects specific to 3D virtual environments. In addition, these dimensions fell short in evaluating technical aspects of computer-aided collaboration. While they offered some metrics concerning information pooling, reaching consensus, task division and shared task alignment, they did not elaborate how such metrics could be measured. Another very similar model was presented in (Burkhardt et al. 2009) and the same shortcomings apply to it.

One noteworthy contribution involving a generic framework, has been offered by Oliveira, Shirmohammadi, and Georganas (1999) who assess the technical aspects of CVE based on four-delay evaluation of display, graphics, simulation and interconnection. Display evaluation is measured by screen delay. Graphics evaluation measures delay caused by the graphics card. Simulation evaluation measures hardware and software processing lag. Interconnections evaluation measures networking and communication delays. However, evaluation of the technical aspects of a CVE is not enough to judge the overall usability and usefulness of CVE.

In other work focusing on evaluation of the behavioural aspects of a CVE, Hrimech and Merienne (2010) presented a generic evaluation framework that can assess different type of VEs. The authors used three measures in their proposed framework: Copresence, Involvement/awareness, Collaborative effort. In spite of its goal, this framework appears limited to the technical aspect of copresence in a CVE. Moreover, this framework did not assess the behavioural aspects in depth, instead the collaborative effort was assessed using a survey of four items. In addition, the evaluation tool was limited to a post session survey of ten items that is not enough to judge the complex nature of CVE.

3 THE PROPOSED EVALUATION FRAMEWORK (THE CORE)

To create an evaluation framework for 3D CVEs, it is important to determine the goal of evaluation, the important factors to evaluate and the suitable evaluation tools to use. A number of studies have been working on different evaluation goals including usability, effectiveness, confidence, interaction and co presence (Tsiatsos and Andreas 2010). Trying to evaluate all possible criteria is almost impossible. Moreover, some of these criteria are contradictory or conflicting, for instance usability criteria may require easy and open access while security criteria may require highly restricted access and a complex method of gaining access to the CVE. Our proposed framework is designed to be generic and suitable for different purposes of CVE and to allow multiple criteria to be evaluated.

Considering all these challenges, we propose an evaluation framework (THE CORE) to assess users' performance, satisfaction and comfort while carrying out a collaborative task. The proposed evaluation framework consists of four levels of evaluation (see Table. 1). 3D CVE application usability sits at the bottom general level and concerns evaluation of the participants' interaction with the whole CVE application. The next layer is a more specific level and seeks to evaluate the participants' interaction with the available tools in the CVE. On a third level of evaluation, companion interaction considers how well the usability of both the application and tools work towards supporting the participants in the collaborative activity. A number of studies stressed the importance of assessing the collaboration outcome as an indicator of a successful CVE (Bruckman and Bandlow 2002). The final level of evaluation includes the collaboration outcome as a result of interaction from the previous level. We claim that our proposed evaluation model shows a *cause and effect* relationship between the VE and its components on one hand, and the effect on collaborative interaction and the outcome of collaboration on the other hand. In the following subsections, the four-layered evaluation framework is briefly introduced. The factors used in each layer, their sources/justification, evaluation criteria and examples are presented later in the results section in Table 2.

3.1 CVE Application

The first level of evaluation is to assess the software application that hosts the CVE. After surveying many research studies that consider (C)VE evaluation, a number of prominent factors were determined (see Table 2). In addition, for each factor some examples of evaluation criteria were provided. These evaluation criteria can be adopted and adapted in accordance with the nature of the target 3D CVE.

3.2 CVE User Interface

Going a step further in evaluation, the second level in the proposed framework is the CVE user interface. This layer aims to assess users' experience with the interface. Understanding the behavior of the learner while using the system helps to determine if a tool is error-prone or hinders the flow of interaction. Evaluating the user interface of 3D CVE has drawn the interest of many studies (e.g. (Xiangyu and Dunston 2006)). Some salient controlling factors to assess the user interface were identified (see Table 2).

3.3 Companion Interaction

In the companion interaction level, the usability of the CVE tool is evaluated. This evaluation is represented by the measurement of companion interaction using these tools. Table 2 shows the two factors to evaluate this level, i.e. involvement and mutual understanding and evaluation criteria. These evaluation criteria include objective measurements such as tracking ratio of agreement/disagreement between users. Moreover, the evaluation criteria includes capturing participants' experience of CVE such as surveying participants' perspective about companion's behaviour as well as the perspective about companions' match/mismatch understanding.

Measured Level	Source	Level of Evaluation	Evaluation Tools	Reason for Tool Selection
Collaborative Outcome	(Biocca et al. 2001), (Bruckman and Bandlow 2002)	Level 4: Outcome	– Surveying performance perception. – Tracking the objective measurements of performance.	Needs both objective evaluation as well as personal perspective
Companion Interaction	(Macdonald 2003), (Kreijns et al. 2003)	Level 3: Interaction	– Surveying effective interaction perception. – Tracking the objective measurements of performance.	Needs both objective evaluation as well as personal perspective
CVE User Interface	(Xiangyu and Dunston 2006)	Level 2: Internal structure	– Tracking participants' interaction with VE tools.	Needs objective evaluation
CVE Application Interaction	(Mikovec et al. 2007)	Level 1: Application	– Tracking participants' interaction with VE application.	Needs objective evaluation

Table 1. Four-layered evaluation framework for CVE (THE CORE)

3.4 Collaboration Outcome

According to an intensive survey of learning CVE (see (Gress et al. 2010)), the majority of assessment strategies rely on the final achievement that is to say the product, rather than the process, of collaboration.

Considering only the final output of 3D CVE tends to narrow down the scope of assessment and ignore important view to what happened during the collaboration. The top level of the proposed CVE evaluation is collaboration outcomes. This level includes two factor the performance while executing the collaborative task and the final achievement. Table 2 shows some examples of evaluation criteria. These evaluation criteria includes both objective measurements as well as participants' perception about to performance and achievement. In order to assess performance, achievement and completion time, and five items for each included in a post session survey.

Measured Level	Factors	Source	Examples of Evaluation Criteria
Collaboration Outcome	Achievement (product)	(Macdonald 2003)	– Survey participants' satisfaction with collaboration achievement. – Track average time to complete the task. – Track Success ratio to complete the collaborative task.
	Performance (process)	(Hrimech and Merienne 2010)	– Survey participants' satisfaction with collaboration performance. – Track the time to complete the collaborative task. – Track/ compare the output of CVE and face-face collaboration.
Companion Interaction	Involvement and co presence	(Wright and Madey 2009) (Heldal et al. 2007)	– Survey the perspective about companion's behaviour. – Survey the interest in future similar CVE.
	Concurrency control	(Bruemmer et al. 2005)	– Track # of turn taking collision.
	Mutual understanding	(Bruemmer et al. 2005) (Spada et al. 2005)	– Survey the perspective about companions' match/mismatch understanding. – Track # of successful interaction to achieve a procedure and ratio of agreement/disagreement
	Communication awareness	(Xiangyu and Dunston 2006)	– Survey participants' attitude toward collaborators' capabilities.
CVE User Interface	Functionality	(Tsiatsos and Andreas 2010)	– Track # of success result of using tools. – Track ratio of tools usage.
	Usability	(Xiangyu and Dunston 2006) (Schroeder et al. 2006)	– Track # of clicks to achieve a procedure. – Track # of plausible navigations/ movements to achieve a procedure.
	Consistency	(Heldal et al. 2007)	– Survey users' perspective about the usability of CVE tools.
CVE Application	Robustness.	(Mikovec et al. 2007)	– Track Crash ratio.
	Usability	(Mikovec et al. 2007)	– Track # of help request during collaborative task.
	Compatibility	(Wright and Madey 2009) (Holmberg, Wunsche, and Tempero 2006)	– Track # of incompatibility errors while installing/running VE application. – List the required hardware and determine whether users are compatible with.

Table 2. THE CORE's levels, factors, source and examples of evaluation criteria

3.5 Data Collection and Analysis

To evaluate the proposed model, suitable methods for data collection are required to measure each level. Gress et al. (2010) conducted an extensive literature search for all articles related to CVE in the learning field, and they found that surveys are the most common and successful evaluation tools to capture participants' responses and experiences. Processing collaboration and discussion data came as a next most successful evaluation method. A more recent evaluation method includes investigating participants' experiences via automatic real-time data capturing (Winne et al. 2006).

Based on the previous findings, two techniques have been chosen to evaluate the proposed model: 1) tracking mechanism, or data logging and 2) survey. These techniques have been found in the literature to be effective means for evaluation, but they have to be adapted to the current purpose. Tracking allows capture of user input, potentially from multiple input devices, and navigation paths allowing recreation and monitoring of what the user actually does. Thus, the tracking mechanism was selected as an objective tool to evaluate the usability of VE application as a whole and the usability of the tools in 3D CVE. However, measurement of companion interaction and collaboration outcomes depend more on the personal and subjective perspective that can supplement the use of the objective tools. To evaluate the CVE thoroughly, the personal perspective should be taken into consideration. After all, a highly usable 3D CVE does not necessarily ensure effective collaborative nor satisfactory outcomes. Survey questions are used as an evaluation tool to capture participants' experience with the CVE. Table 1 shows the evaluation tools used for each level, the source justifying each tool along with the reason behind selection of this tool.

4 CASE STUDY

In order to test the proposed evaluation framework, a 3D CVE was implemented and an empirical experiment was conducted. The experiment involved fifty-five (55) second-year undergraduate students. A post session survey was used to gather participants' responses about the interaction with the IVA and about collaboration outcome. In addition to a survey, automatic tracking was used to capture all transactions and interactions in log files.

Figure 1. A snapshot from the case study

According to the implemented 3D CVE, the human participants have to collaborate with a virtual agent, see Figure 1, to reach their target. The collaborative task is to cross a sequence of obstacles using a pair of tools from a toolbox that contains twelve tools to reach a research laboratory. For each obstacle, both the human and IVA will take alternative turns to decide which tool is the best choice to cross this obstacle. While participants and IVAs collaborate, they have to exchange messages to organize their teamwork and use some tools to achieve their target goal.

5 RESULTS

After the completion of the experiment, data was combined and the proposed evaluation framework was applied. Descriptive statistics were used to analyse the quantitative data. Some of the results are presented in the following subsection.

5.1 CVE Application

The CVE used to test the evaluation framework is a web-based application; no compatibility error occurred in running the system (Windows, Mac OS and Linux). The required input/output devices are traditional keyboard, mouse and monitor. Using traditional input/output devices adds to CVE ease of use and no issues were found. Logging crash report and calculating ratio of crash execution to successful execution demonstrated the robustness of the application. The current case study shows high level of compatibility and robustness. However, as the CVE was designed using the Unity 3D game engine a plug-in needed installation beforehand, thus, this requirement may decrease the assessment of usability factor.

5.2 CVE User Interface

To evaluate the user interface of CVE, a number of factors were proposed. These factors cover the design and the functionality of the CVE. Table 2 presents factors for evaluating the user interface and examples of evaluation criteria. Table 3 provides the usability and functionality ratios for the tools in the implemented CVE. To reach successfully the collaborative goal, the usability ratio for each tool should contribute on average 8.33% (100% ÷ 12). The result show that the usability ratio of the majority of the tools (11 out of 12) is close to average (M= 7%, SD=3.19%). The results shows that usability ratio of "rope" tool was above average. Exceeding the average usability ratio suggests that the participants tried to use this tool where it is not needed to be utilized. Regarding the functionality ratio, the result show that the majority of the tools were used successfully (over 90%) to achieve the function it was designed for. Nevertheless, the tool rope shows low functionality ratio. The abnormal ratio of usability and functionality in some user interface tools necessitates a revision to these tools to make sure use of appropriate usability and proper functionality.

Tools	pruning shears	bush hook	hammer	chisel	ladder	rope	match sticks	match box	screw drive	nipper	shovel	mattock
Usability Ratio	4.39%	12.79%	2.51%	3.13%	11.70%	22.43%	6.90%	6.13%	7.76%	6.27%	7.52%	8.46%
Functionality Ratio	94.65%	93.42%	97.80%	95.21%	89.88%	79.67%	92.88%	89.13%	92.44%	89.15%	95.71%	90.36%

Table 3. Evaluation of the user interface of the case study

5.3 Companion Interaction

The results (see Figure 2) showed that 80.00% of participants agreed and strongly agreed on the development of mutual understanding with their collaborator. Furthermore, the results show that 78.95% of participants answered with agree and strongly agree that the CVE encouraged them to involve in the collaborative activity, while 81.82% were satisfied with partner's communication capabilities during the collaborative session. The case study followed sequential turn taking. Participant cannot begin his/her turn, unless the other partner finish his/her role; hence, there was no concurrency problems raised. There is no exact threshold to judge the level of satisfaction with collaboration; however, it is a reference point for CVE designer to determine whether to accept the result of current evaluation or work on improving the factors of companion interaction.

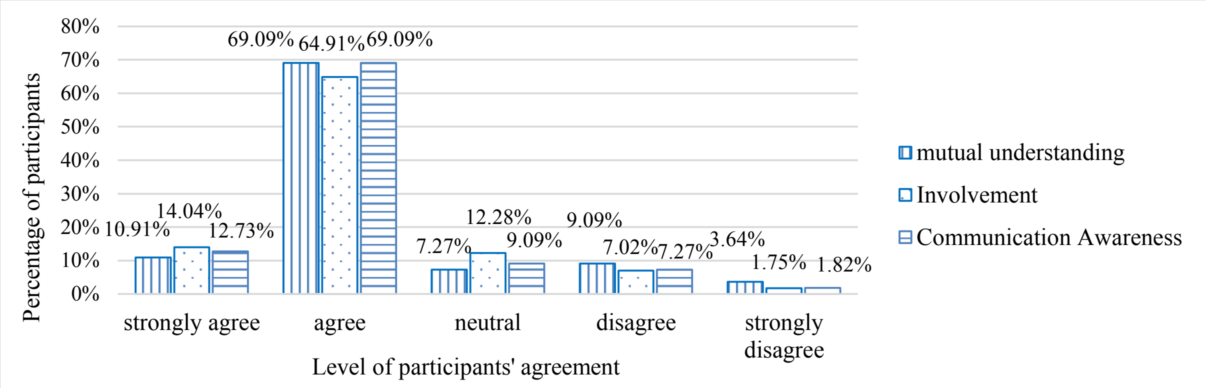

Figure 2. *Perception of CVE participants' mutual understanding, involvement, and communication awareness*

5.4 Collaboration Outcome

The result (see Figure 3) showed that 81.82% and 72.73% of the participants reported that they were satisfied (agreed and strongly agreed) with the achievement and the performance respectively, while achieving the goal of CVE. In addition, the result showed that 76.36% of the participants were satisfied with time management they planned with their collaborators. As the ratio of collaboration performance tends to be less than the average of other factors, this could be a reference to CVE designer that the performance in this case study needs further improvement.

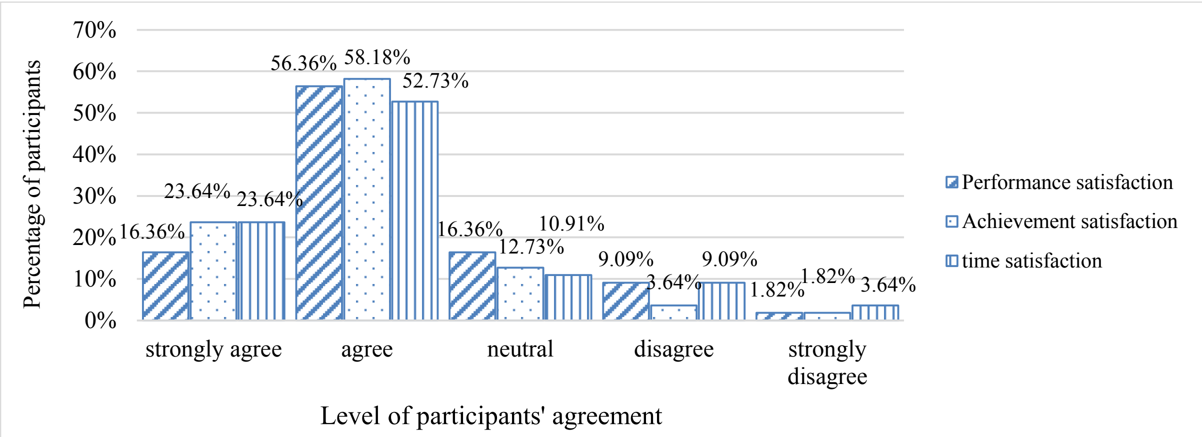

Figure 3. *Perception of CVE participants' performance, achievement and time management*

6 CONCLUSION AND FUTURE WORK

This paper presented an evaluation framework for 3D CVEs. To address the gap in current studies, the proposed framework was designed as a four-layered structure that aimed to be: first, generic to evaluate different applications of CVEs; second, thorough to evaluate the prominent aspects of CVEs.

As future work, the proposed evaluation framework will be applied to different types of CVEs, such as CVEs for learning, business, entertainment or education. Applying the framework to different CVEs will help to identify new factors and modify the current factors in each layer. Another future research plan is to include more examples of evaluation criteria that match various CVE design. Possible extensive application of the framework to other reported CVE studies may also lead to further criteria and examples.

References

Biocca, F., C. Harms, and J. Gregg (2001). The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. In *Fourth Annual International Workshop on Presence*. Temple University, Philadelphia, PA.

Bishop, Ian D. and Christian Stock (2010). Using collaborative virtual environments to plan wind energy installations. *Renewable Energy* 35, no. 10: 2348-2355.

Bowman, Doug A., Donald B. Johnson, and Larry F. Hodges (1999). Testbed evaluation of virtual environment interaction techniques. Paper presented at Proceedings of the ACM symposium on Virtual reality software and technology, London, United Kingdom.

Bruckman, A. and A. Bandlow (2002). *Hci for kids: The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications*. Ed J. Jacko and A. Sears: Lawrence Erlbaum and Associates.

Bruemmer, D. J., D. A. Few, R. L. Boring, J. L. Marble, M. C. Walton, and C. W. Nielsen (2005). Shared understanding for collaborative control. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans* 35, no. 4: 494-504.

Burkhardt, Jean-Marie, Françoise Détienne, Anne-Marie Hébert, and Laurence Perron (2009). Assessing the "quality of collaboration" in technology-mediated design situations with several dimensions. In *Human-computer interaction – interact 2009*, ed. Tom Gross, Jan Gulliksen, Paula Kotzé, Lars Oestreicher, Philippe Palanque, RaquelOliveira Prates and Marco Winckler, 5727:157-160: Springer Berlin Heidelberg.

Caballé, Santi, Thanasis Daradoumis, Fatos Xhafa, and Angel Juan (2011). Providing effective feedback, monitoring and evaluation to on-line collaborative learning discussions. *Computers in Human Behavior* 27, no. 4: 1372-1381.

Chee, Yam San and Chit Meng Hooi (2002). C-visions: Socialized learning through collaborative, virtual, interactive simulations. Paper presented at Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community, Boulder, Colorado.

Giraldo, Faber, Àngela María, Jiménez Rojas, Pedro Esteban, and Helmuth Trefftz (2007). Collaborative virtual environments for teaching physics. In *Innovations in e-learning, instruction technology, assessment, and engineering education*, ed. Magued Iskander:89-93: Springer Netherlands.

Greenberg, Saul and Bill Buxton (2008). Usability evaluation considered harmful (some of the time). Paper presented at Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Florence, Italy.

Gress, Carmen L. Z., Meghann Fior, Allyson F. Hadwin, and Philip H. Winne (2010). Measurement and assessment in computer-supported collaborative learning. *Computers in Human Behavior* 26, no. 5: 806-814.

Heldal, Ilona, David Roberts, Lars Bråthe, and Robin Wolff (2007). Presence, creativity and collaborative work in virtual environments. In *Human-computer interaction. Interaction design and usability*, ed. JulieA Jacko, 4550:802-811: Springer Berlin Heidelberg.

Holmberg, N., B. Wunsche, and E. Tempero (2006). A framework for interactive web-based visualization. In *the 7th Australasian User interface conference (AUIC '06)*:137–144. Darlinghurst, Australia: Australian Computer Society, Inc.

Hrimech, Hamid and Frédéric Merienne (2010). Interaction and evaluation tools for collaborative virtual environment. *International Journal on Interactive Design and Manufacturing (IJIDeM)* 4, no. 3: 149-156.

Kreijns, K., P. A. Kirschner, and W. Jochems (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: A review of the research. *Computers in Human Behavior* 19: 335-353.

Lorenzo, Carlos-Miguel, Miguel Ángel Sicilia, and Salvador Sánchez (2012). Studying the effectiveness of multi-user immersive environments for collaborative evaluation tasks. *Computers & Education* 59, no. 4: 1361-1376.

Macdonald, Janet (2003). Assessing online collaborative learning: Process and product. *Computers & Education* 40, no. 4: 377-391.

Mikovec, Z., I. Maly, P. Slavik, and J. Curin (2007). Visualization of users' activities in a specific environment. In *the 2007 Winter Simulation Conference*, ed. S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew and R. R. Barton:738-746.

Oliveira, Jauvane C. de, Shervin Shirmohammadi, and Nicolas D. Georganas (1999). Collaborative virtual environment standards: A performance evaluation. Paper presented at Proceedings of the 3rd International Workshop on Distributed Interactive Simulation and Real-Time Applications.

Reis, R., B. Fonseca, and P. Escudeiro (2013). Educational collaborative virtual environments: Evaluation model. In *IEEE 13th International Conference on Advanced Learning Technologies (ICALT'13)*:215-216.

Schroeder, Ralph, Ilona Heldal, and Jolanda Tromp (2006). The usability of collaborative virtual environments and methods for the analysis of interaction. *Journal of Presence* 15, no. 6: 655-667.

Snowdon, Dave, ElizabethF Churchill, and AlanJ Munro (2001). Collaborative virtual environments: Digital spaces and places for cscw: An introduction. In *Collaborative virtual environments*, ed. ElizabethF Churchill, DavidN Snowdon and AlanJ Munro:3-17: Springer London.

Spada, Hans, Anne Meier, Nikol Rummel, and Sabine Hauser (2005). A new method to assess the quality of collaborative process in cscl. Paper presented at Proceedings of the 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years!, Taipei, Taiwan.

Tsiatsos, Thrasyvoulos and Konstantinidis Andreas (2010). Evaluation framework for collaborative educational virtual environments. *Journal of Educational Technology & Society* 13, no. 2: 65-77.

Wang, Qiyun (2009). Design and evaluation of a collaborative learning environment. *Computers & Education* 53, no. 1: 1138–1146.

Whalen, T. E., S. Noel, and J. Stewart (2003). Measuring the human side of virtual reality. In *IEEE International Symposium on Virtual Environments, Human-Computer Interfaces and Measurement Systems (VECIMS '03)*:8-12.

Winne, P. H., J. C. Nesbit, V. Kumar, A. F. Hadwin, S. P. Lajoie, R. A. Azevedo, and NANCY E. Perry (2006). Supporting self-regulated learning with gstudy software: The learning kit project. *Technology, Instruction, Cognition and Learning* 3: 105-113.

Wright, Timothy E. and Greg Madey (2009). A survey of technologies for building collaborative virtual environments. *The International Journal of Virtual Reality* 8, no. 1: 53-66.

Xiangyu, Wang and P. S. Dunston (2006). Usability evaluation of a mixed reality collaborative tool for design review. In *Computer Graphics, Imaging and Visualisation, 2006 International Conference on*:448-451.

Zhenyu, Yang, Yu Bin, Wu Wanmin, K. Nahrstedt, R. Diankov, and R. Bajscy (2006). A study of collaborative dancing in tele-immersive environments. In *Eighth IEEE International Symposium on Multimedia (ISM'06)*:177-184.